Der Berner Stichprobenplan. Ein Vorschlag für eine effiziente Klumpenstichprobe am Beispiel der Schweiz[1]

Ben Jann

ETH Zürich

Version vom 9. März 2006

Erscheint voraussichtlich im Reader zur Tagung "Stichprobenqualität in Bevölkerungsumfragen" der ASI und der DGS-Methodensektion, Berlin, 14.-15. Oktober 2005.

Die meisten Bevölkerungsumfragen in der Schweiz beruhen auf Stichproben, die aus dem Telefonregister gezogen wurden. Da solche Stichproben bezüglich der Abdeckung der Grundgesamtheit als problematisch anzusehen sind, diskutiere ich hier ein alternatives, von Fritschi et al. (1976) entwickeltes Stichprobenverfahren, bei dem in einem ersten Schritt Gemeinden ausgewählt und dann die Adressen der Zielpersonen über die Einwohnerregister gezogen werden. Das Verfahren ist dabei so angelegt, dass die Stichprobe ähnlich wie bei einer einfachen Klumpenstichprobe auf eine relativ geringe Anzahl Gemeinden verdichtet wird, die Stichprobe aber trotzdem eine möglichst hohe statistische Effizienz beibehält. Eine Analyse der theoretischen Eigenschaften des Berner Stichprobenplans zeigt, dass das ursprüngliche Verfahren zu leicht verzerrten Stichproben führt. Eine korrigierende Modifikation des Verfahrens wird vorgeschlagen. Zudem wird ein alternativer Ansatz mit verbesserten Eigenschaften, die *ex ante* geteilte Stichprobe, vorgestellt. Mit Hilfe einer Simulationsstudie werden sodann die Vorzüge des Berner Stichprobenplans gegenüber einer einfachen Klumpenstichprobe illustriert.

1 Einleitung

Wie auch in anderen Ländern ohne zentrales Einwohnerregister stellt in der Schweiz die Ziehung einer repräsentativen Bevölkerungsstichprobe die empirische Sozialforschung vor schwerwiegende Probleme. Ohne Urliste der Mitglieder der zu untersuchenden Population lässt sich naturgemäß nur schwer eine unverzerrte Zufallsstichprobe aus derselben ziehen. Zudem ist es, wenn man über ein ansprechendes und praktikables Stichprobenverfahren verfügt, auch nicht immer ganz einfach festzustellen, ob das Verfahren tatsächlich systematisch unverzerrte Resultate liefert. Die Erstellung von qualitativ hochwertigen Stichproben kann also sehr zeitaufwändig sein und verursacht unter Umständen erhebliche Kosten. Somit ist wohl nicht zuletzt aufgrund beschränkter Forschungsbudgets damit zu rechnen, dass ein großer Teil der empirischen Umfrageforschung in der Schweiz auf zumindest zweifelhaften Stichproben beruht.

Die einfachste und wahrscheinlich am Häufigsten angewandte Stichprobenmethode in der Schweiz ist die Ziehung einer Haushaltsstichprobe aus dem Telefonverzeichnis (vgl. etwa Jann 2001; für Deutschland Schnell 1997), die dann in einem zweiten Schritt nach der Erfassung der Haushaltsstrukturen durch Ziehung von einzelnen Zielpersonen in

[1] In Gedenken an Herbert Iff († 18. April 1998).

den Haushalten in eine Personenstichprobe überführt wird. Die Probleme dieser Methode sind bekannt: Erstens gibt es einen gewichtigen Anteil der Bevölkerung, der auf diesem Weg nicht erreicht werden kann, weil kein oder nur ein nicht-registrierter Telefonanschluss vorliegt (zu etwas älteren Schätzungen dieses Anteils siehe zum Beispiel die Studien von Schmugge und Grau 1998, 2000; Experten gehen heute von einem Anteil nicht über das Telefonverzeichnis erreichbarer Personen von 10 bis 15 Prozent aus). In der Schweiz gibt es nach wie vor Haushalte ohne Telefonanschluss und seit 1998 kann die sonst automatische Aufnahme eines Festnetzanschlusses ins öffentliche Telefonverzeichnis verweigert werden. Weiterhin werden Mobiltelefonanschlüsse nur auf speziellen Wunsch hin in das Verzeichnis aufgenommen, was eine entsprechend geringe Registrierungsquote zur Folge hat. Zweitens handelt es sich bei der Ziehung aus dem Telefonregister wie angesprochen zumeist um ein zweistufiges Verfahren, bei dem innerhalb der Haushalte auf der zweiten Stufe eine Personenauswahl getroffen wird. Man kann sich hierbei leicht zusätzliche Verzerrungen einhandeln, wenn die Auswahl der Zielperson im Haushalt zum Beispiel nach dem Schwedenschlüssel oder mit der Geburtstagsmethode durch die Haushalte selbst vorgenommen wird (dies gilt allerdings nicht für CATI-Umfragen, bei denen die Auswahl der Zielperson normalerweise per Computer nach telefonischer Erfassung der Haushaltsstruktur erfolgt). Zudem führt die Zweistufigkeit des Verfahrens meistens zu einer Stichprobe mit nicht-identischen Auswahlwahrscheinlichkeiten, was tendenziell zu geringerer Stichprobeneffizienz führt und die Unannehmlichkeiten einer Gewichtung bei der Datenanalyse nach sich zieht. Weitere Probleme von Telefonstichproben betreffen zum Beispiel das Vorhandensein von Mehrfacheinträgen im Telefonregister beziehungsweise von unterschiedlichen Einträgen, die auf die gleichen Haushalte verweisen, und den damit verbundenen Schwierigkeiten, die Auswahlwahrscheinlichkeiten einzelner Personen genau zu bestimmen.

Alternativen zu den Stichproben aus dem Telefonregister gibt es wenige, so dass zum Beispiel auch das Schweizerische Bundesamt für Statistik für etliche Großbefragungen wie etwa die jährlich durchgeführte Schweizerische Arbeitskräfteerhebung (SAKE) darauf zurückgreift. Auch stecken Random-Digit-Dialing-Verfahren (RDD), mit denen man zumindest die nicht eingetragenen Haushalte erreichen könnte, in der Schweiz noch in den Kinderschuhen. Erste Gehversuche werden zurzeit vom LINK-Institut unternommen und zudem scheint die BIK Aschpurwis + Behrens GmbH aus Hamburg neuerdings RDD-Stichproben für die Schweiz anzubieten, die sich an das Verfahren von Gabler und Häder anlehnen (1997; vgl. auch den Übersichtsartikel von Häder und Glemser 2006). Ein Problem ist aber natürlich auch bei RDD-Verfahren, dass ein Teil der Bevölkerung so nicht erreichbar ist und der Schritt vom Telefonanschluss zur effektiv zu interviewenden Zielperson steinig sein kann.

Telefonstichproben (und RDD-Stichproben, wenn ein entsprechendes Verfahren einsatzbereit ist) mögen also zwar verhältnismäßig einfach und preiswert zu erlangen sein, sie sind aber aus den angesprochen Gründen methodisch nicht wirklich befriedigend. Eine bessere Datenquelle zur Ziehung von qualitativ hochwertigen Stichproben wären da ohne Zweifel die Einwohnerregister der Gemeinden. Die Register der Gemeinden, obwohl zurzeit noch in recht heterogener Form bezüglich der Art der Datenerfassung und der enthaltenen Information vorliegend, geben sicherlich das aktuellste und vollständigste Abbild der Bevölkerung der Schweiz.[2] Man beachte, dass die Stichprobenziehung über die Gemeinderegister unter Umständen sogar den Einbezug der Anstaltsbevölkerung in die Stichprobe erlaubt (vgl. zum Thema Schnell 1991). Nicht abgedeckt werden einzig einige Randgruppen wie illegal eingewanderte Personen, die sich der Registrierung durch die Behörden bewusst entziehen. Geringfügige Unschärfen ergeben sich zudem aufgrund von Mobilitätsprozessen. Insgesamt ist aber an den Einwohnerregisterdaten eigentlich nur problematisch, dass die Daten nicht in zentralisierter Form vorliegen. Entschließt man sich, eine Stichprobe mit Hilfe der Register zu ziehen, muss man in einem ersten Schritt also ein Verfahren zur Auswahl von Gemeinden anwenden. In einem zweiten Schritt wird dann mit den Gemeindeverwaltungen Kontakt aufgenommen, um die Zielpersonen der Stichprobe zu bestimmen.

Ein Gemeinde-Auswahlverfahren kann sein, mit Hilfe des vom Bundesamt für Statistik geführten Gemeindeverzeichnisses, das unter anderem Angaben zur Anzahl Einwohner in den Gemeinden enthält, eine künstliche Liste der Schweizer Bevölkerung zu erstellen, und aus dieser Liste mit einem Zufallsalgorithmus eine einfache Wahrscheinlichkeitsauswahl zu ziehen. Gemeinden, die mit mehr als einem Treffer in der fiktiven Stichprobe enthalten sind, werden dann zwecks Ziehung einer entsprechenden Anzahl realer Zielpersonen aus deren Einwohnerregister angeschrieben. Der Nachteil dieses Verfahrens ist, dass es zu einer Stichprobe mit sehr feiner geographischer Granularität führt. Das heißt, in einer so gezogenen Stichprobe sind sehr viele Gemeinden enthalten und in einem großen Teil dieser Gemeinden muss vielleicht nur gerade je ein Interview durchgeführt werden. Eine Folge davon ist, dass die Adressbeschaffung sehr zeitaufwändig und auch kostspielig wird. Zudem entstehen etwa bei Face-to-Face-Interviews hohe Reisespesen.

Eine Lösung zur Verminderung der geographischen Streuung und somit der Reduktion der Kosten ist die Ziehung einer Klumpenstichprobe, bei der zum Beispiel Gemeinden mit einem PPS-Verfahren (*Probability Proportional to Size*) gezogen werden, und dann pro gezogene Gemeinde eine fixe Anzahl Interviews durchgeführt wird.[3] Die Anzahl in der Stichprobe enthaltenen Gemeinden kann so im Vergleich zu einer einfachen Wahrscheinlichkeitsauswahl stark reduziert werden. Allerdings leidet aber auch die statistische Effizienz, das heißt, die Schätzung von Populationsparametern auf Grundlage solcher Klumpenstichproben ist unter Umständen mit einer deutlich erhöhten Unsicherheit behaftet.

Vor dem Hintergrund, dass eine einfache Wahrscheinlichkeitsauswahl zwar gute statistische Eigenschaften aufweist, aber nur aufwändig zu erlangen ist, und umgekehrt bei einer kostengünstigen Klumpenstichprobe die statistische Effizienz zu wünschen

---

[2] Auf Bundesebene gibt es konkrete Bestrebungen zur Harmonisierung der Einwohnerregister insbesondere auch zum Zwecke bevölkerungsstatistischer Erhebungen. So wird in den Räten zurzeit ein neues Registerharmonisierungsgesetz (RHG) verhandelt.

[3] Beziehungsweise in großen Gemeinden unter Umständen ein Vielfaches dieser fixen Anzahl. Bei sehr schiefen Verteilungen wie der Gemeindgrößenverteilung macht nur ein PPS-Verfahren *mit Zurücklegen* Sinn (einzelne Gemeinden hätten in einem Verfahren ohne Zurücklegen Auswahlwahrscheinlichkeit größer eins), so dass eine Gemeinde mehrmals in die PPS-Stichprobe gelangen kann. In solchen Gemeinden ist ein entsprechendes Vielfaches an Zielpersonen zu ziehen.

übrig lässt, haben Fritschi, Meyer und Schweizer (1976) einen Vorschlag für eine „geteilte“ Zufallsstichprobe ausgearbeitet. Das Verfahren hat sich in Anlehnung an die Herkunft der Autoren als der „Berner Stichprobenplan“ in der Literatur niedergeschlagen. Zwar wurde der Stichprobenplan in der Schweiz für eine Reihe von Studien verwendet (die prominenteste Arbeit, bei der auf das Verfahren zurückgegriffen wurde, ist die Armutsstudie von Leu et al. 1997; vgl. weiterhin z.B. Meyer et al. 1982 oder Wydler et al. 1996). Da es sich beim Berner Stichprobenplan aber um einen viel versprechenden Ansatz handelt, dessen Vorteile gegenüber den üblichen Telefonregisterstichproben meiner Meinung nach bisher zu wenig genutzt werden, möchte ich nachfolgend einige Überlegungen zu diesem Verfahren präsentieren.

Im nächsten Abschnitt wird das von Fritschi et al. (1976) vorgeschlagene Verfahren genauer beschrieben. Es wird unter anderem gezeigt, dass das Verfahren zu leicht verzerrten Stichproben führt, und es wird ein korrigiertes Verfahren entwickelt, das diese Verzerrungen ausgleicht. Der dritte Abschnitt befasst sich mit einer nahe liegenden Vereinfachung des Verfahrens und es werden die Eigenschaften des Berner Stichprobenplanes anhand von Simulationsergebnissen illustriert. Der vierte Abschnitt fasst die Ergebnisse zusammen.

2 Das „geteilte“ Stichprobenverfahren von Fritschi, Meyer und Schweizer

Der Ausgangspunkt des Klumpungsverfahrens nach Fritschi et al. (1976) ist die Überlegung, dass eine Klumpung der Stichprobe für relativ große Gemeinden wenig Sinn macht, weil auf diese Gemeinden auch bei einer einfachen Wahrscheinlichkeitsauswahl eine hinreichend große Anzahl Stichprobenmitglieder entfällt. Durch die Klumpung würde man in diesem Teil der Population unnötigerweise eine Menge Präzision verschenken. Das Problem sind vielmehr die vielen kleinen Gemeinden, in denen bei einer einfachen Wahrscheinlichkeitsauswahl jeweils nur eine minimale Anzahl Interviews durchgeführt werden muss. Fritschi et al. (1976) schlagen deshalb das folgende Verfahren zur Erstellung einer „geteilten“ Zufallsstichprobe vor: In einem ersten Schritt ziehe man ausgehend von einem Gemeindeverzeichnis mit Angaben zu den Einwohnerzahlen (bzw. der Anzahl Stimm- und Wahlberechtigten im Anwendungsfall von Fritschi et al.) eine (hypothetische) Stichprobe aus der durch das Gemeindeverzeichnis definierten Population. Die Ziehung erfolgt nach den Regeln einer einfachen Wahrscheinlichkeitsauswahl mit konstanter Auswahlwahrscheinlichkeit $p = n / N$, wobei n dem gewünschten Umfang der Stichprobe und N der Populationsgröße entspricht. Für jede Gemeinde wird dann die Anzahl „Treffer“, also die Anzahl hypothetischer Stichprobenmitglieder, die auf diese Gemeinde entfallen sind, gezählt. Aus der durch die Gemeinden mit weniger als k Treffern definierten Teilpopulation wird dann in einem zweiten Schritt eine Stichprobe mit reduzierter Trefferwahrscheinlichkeit p / k gezogen. Die Endstichprobe setzt sich schließlich zusammen aus einer der Anzahl Treffer entsprechen Anzahl Zielpersonen in den Gemeinden mit k oder mehr Treffern im ersten Schritt und jeweils k Zielpersonen in den Gemeinden mit einem Treffer im zweiten Schritt (bzw. ein der Anzahl Treffer entsprechendes Vielfaches von k, wenn eine Gemeinde im zweiten Schritt mehr als einen Treffer verzeichnet). Die Bestimmung der effektiven Zielpersonen erfolgt dann in der Praxis am besten über die Einwohnerregister der in der Stichprobe enthaltenen Gemeinden. Alternativ wäre auch ein Random-Route-Verfahren denkbar.

Der Ansatz von Fritschi et al. (1976) ist einleuchtend: Zuerst eine einfache Wahrscheinlichkeitsauswahl durchführen, dann denjenigen Teil der Stichprobe, in dem nur wenige Treffer pro Gemeinde vorliegen, mit einer Klumpenstichprobe ersetzen. Anders als von Fritschi et al. vermeintlich bewiesen, handelt es sich aber nicht um eine Methode, bei der die *a priori* Auswahlwahrscheinlichkeiten in allen Gemeinden gleich sind, und die Verteilungen von mit der Gemeindegröße zusammenhängenden Merkmalen wird verzerrt wiedergegeben. Dies soll nun dargelegt werden.

Sei P_{ij} die Wahrscheinlichkeit, dass Person *i* aus Gemeinde *j* in die Stichprobe gelangt. Bei einer einfachen Wahrscheinlichkeitsauswahl ist P_{ij} für alle Personen aus der Grundgesamtheit gleich, also

$$P_{ij} = p = \frac{n}{N} \quad \text{für alle } i, j$$

wenn *n* die Stichprobenumfang und *N* die Populationsgröße ist. Stichproben, die diese Eigenschaft besitzen, werden manchmal als EPSEM-Stichproben bezeichnet (*Equal Probability of Selection Method*, Babbie 1979, S. 330). Zwar ist es möglich, auch aus Stichproben mit nicht-identischen Auswahlwahrscheinlichkeiten auf die Grundgesamtheit zu schließen, sofern die Auswahlwahrscheinlichkeiten bekannt sind. Im Allgemeinen sind jedoch EPSEM-Stichproben aus theoretischen (mehr statistische Effizienz) sowie praktischen Gründen (es werden weniger komplexe Schätzer benötigt) vorzuziehen (eine Ausnahme mögen je nach Eigenschaften des Untersuchungsgegenstandes disproportional geschichtete Stichproben bilden). Im vorliegenden Fall eines gemeindebasierten Stichprobenverfahrens sollte insbesondere darauf geachtet werden, dass die Auswahlwahrscheinlichkeiten nicht mit der Gemeindegröße zusammenhängen (oder aber, dass der Zusammenhang genau bekannt ist und mit entsprechenden Gewichten bei der Datenanalyse ausgeglichen werden kann).

Wie erläutert wird beim zweistufigen Berner Stichprobenplan in einem ersten Schritt eine einfache Wahrscheinlichkeitsauswahl mit Auswahlwahrscheinlichkeit $p = n/N$ durchgeführt. Es wird sodann die Anzahl „Treffer“, also die Anzahl ausgewählter Personen pro Gemeinde gezählt und in einem zweiten Schritt unter den Gemeinden, die eine kritische Anzahl Treffer *k* nicht erreicht haben, eine Klumpenstrichprobe gezogen. Die Stichprobe setzt sich schließlich zusammen aus den im ersten Schritt gewählten Personen in den Gemeinden mit mehr als *k* Treffern und den Klumpen zu je *k* Personen aus dem zweiten Schritt. Sei *X* eine Zufallsvariable der Anzahl Treffer pro Gemeinde im ersten Schritt und *Y* eine von *X* unabhängige Zufallsvariable der Anzahl Klumpen pro Gemeinde im zweiten Schritt. Die Auswahlwahrscheinlichkeit von Person *i* aus Gemeinde *j* kann dann geschrieben werden als

$$P_{ij} = P\left(X_j \geq k\right) \frac{E\left(X_j \mid X_j \geq k\right)}{N_j} + P\left(X_j < k\right) \frac{E\left(Y_j \mid X_j < k\right) \cdot k}{N_j} \tag{1}$$

wobei $P\left(X_j \geq k\right) = 1 - P\left(X_j < k\right)$ der Wahrscheinlichkeit entspricht, dass in Gemeinde *j* im ersten Schritt *k* oder mehr Treffer erreicht werden. N_j entspricht der Anzahl Personen in Gemeinde *j* und $E\left(.\right)$ symbolisiert den Erwartungswert.

Wird mit Zurücklegen gezogen, können *X* und *Y* als binomialverteilte Zufallsvariablen aufgefasst werden, also

$$X_j \sim B\left(N_j, p\right) \quad \text{und} \quad Y_j \sim B\left(N_j, \pi_j\right)$$

mit *p* und π_j als den Trefferwahrscheinlichkeiten. Es gilt insbesondere

$$E\left(X_j\right) = p \cdot N_j \quad \text{und} \quad E\left(Y_j\right) = \pi_j N_j$$

Aus der Unabhängigkeit von *X* und *Y* folgt zudem $E(Y_j \mid X_j < k) = E(Y_j)$. Wenn der Berner Stichprobenplan nun EPSEM-Charakter haben soll, dann muss $P_{ij} = p = n/N$ gelten (für alle *i* und *j*). Durch Einsetzen in (1) erhält man nach einigen Umformungen jedoch

$$\pi_j = \frac{\dfrac{n}{N} - P\left(X_j \geq k\right) \dfrac{E\left(X_j \mid X_j \geq k\right)}{N_j}}{k \cdot \left[1 - P\left(X_j \geq k\right)\right]} \tag{2}$$

Auch wenn $P\left(X_j \geq k\right)$ und $E\left(X_j \mid X_j \geq k\right)$ weiter aufgelöst werden (siehe Anhang), lässt sich N_j nicht aus Gleichung (2) eliminieren. Das heißt, π_j hängt von der Gemeindegröße ab und ist nicht konstant, was in Widerspruch zu den Ausführungen von Fritschi et al. (1976) steht, die von einer konstanten Auswahlwahrscheinlichkeit $\pi = p / k$ ausgehen. Es ist also zu erwarten, dass der Berner Stichprobenplan nach Fritschi et al., der im ersten wie auch im zweiten Schritt jeweils eine konstante Auswahlwahrscheinlichkeiten vorsieht, zu verzerrten Ergebnissen führt. Die Verzerrung kann indes korrigiert werden, indem im zweiten Schritt variable Auswahlwahrscheinlichkeiten nach Formel (2) eingesetzt werden.[4]

Die Eigenschaften der Verzerrung im unkorrigierten Berner Stichprobenplan können mit Hilfe einer Simulation veranschaulicht werden. Als Grundlage für die Simulation verwende ich das Gemeindeverzeichnis 2002 der Schweiz, das Angaben zu den Einwohnerzahlen aller politischen Gemeinden enthält (Bundesamt für Statistik 2002). Gezogen wird jeweils eine Stichprobe im Umfang von $n = 3750$ Personen, die

[4] Ein alternatives Korrekturverfahren mit einem ähnlichen Effekt bestünde darin, die Auswahlwahrscheinlichkeiten im ersten Schritt variabel zu gestalten, dafür aber im zweiten Schritt konstant zu halten (die erforderlichen Größen lassen sich ebenfalls aus den hier präsentierten Formeln ableiten). Eine weitere Korrekturstrategie wäre, die Stichprobe mit dem unkorrigierten Verfahren zu ziehen, jedoch bei der Datenanalyse aus Formel (1) ableitbare Design-Gewichte einzusetzen.

minimale Anzahl Zielpersonen pro Gemeinde ist auf $k = 10$ festgelegt (dies entspricht bezüglich der Anzahl in die Stichprobe gelangender Gemeinden in etwa den im Anwendungsbeispiel von Fritschi et al. verwendeten Parametern; Fritschi et al. legten ihrem Verfahren die etwas kleinere Population der Wahlberechtigten zu Grunde). Tabelle 1 zeigt die Simulationsergebnisse über 10’000 Runden. Angegeben ist einerseits die faktische Gemeindegrößenverteilung in der Population und andererseits die aus den Stichproben resultierende gemittelte Gemeindegrößenverteilung für verschiedene Ziehungsverfahren. Man erkennt, dass bei der ursprünglichen Version des Berner Stichprobenplans die meisten Gemeindegrößenkategorien leicht unterrepräsentiert sind. Zum Beispiel leben 12.06 Prozent der Bevölkerung in Gemeinden mit 2000 bis 3499 Einwohnern, in den Stichproben auf Grundlage des Berner Stichprobenplans sind es im Schnitt aber nur 11.71 Prozent. Eine Kategorie jedoch, Gemeindegröße 10’000 bis 49’999, ist deutlich übervertreten. Ähnlich wie bei einer einfachen Wahrscheinlichkeitsauswahl (letzte Spalte) treten diese Verzerrungen beim „korrigierten“ Berner Stichprobenplan, bei dem die Auswahlwahrscheinlichkeiten im zweiten Schritt nach Formel (2) bestimmt wurden, jedoch nicht auf.

[Tabelle 1 etwa hier.]

Ein eindrückliches Bild der Natur der Verzerrung im unkorrigierten Berner Stichprobenplan liefert Abbildung 1. Dargestellt ist die relative Verteilung der gemittelten Gemeindegrößenverteilung in den Stichproben und der Gemeindegrößenverteilung in der Population (bzw. genauer: für jede Gemeinde ist das Verhältnis ihres durchschnittlichen Anteils an der Stichprobe und ihres tatsächlichen Anteils an der Gesamtbevölkerung abgebildet). Man sieht sehr deutlich, dass Personen aus Gemeinden mit ca. 10’000 bis 30’000 Einwohnern zu häufig in die Stichprobe gelangen. Natürlich sind dies gerade diejenigen Gemeinden, die an der „Grenze“ zur Klumpenstichprobe liegen. Das sind Gemeinden mit einer durchschnittlich erwarteten Trefferzahl nahe der Mindestklumpengröße *k*, die also je nach Ausgang der einfachen Wahrscheinlichkeitsauswahl im ersten Schritt manchmal in das Klumpenauswahlverfahren im zweiten Schritt gelangen und manchmal nicht. Beim korrigierten Berner Stichprobenplan mit Auswahlwahrscheinlichkeiten nach Formel (2) wird die Verzerrung, wie man erkennen kann, erfolgreich geglättet.[5]

[Abbildung 1 etwa hier.]

[5] Zur Verteidigung von Fritschi et al. (1976) ist anzumerken, dass die hier recht bedeutend erscheinende Verzerrung in der von Fritschi et al. gewählten praktischen Umsetzung ihres Stichprobenverfahrens deutlich geringer war. Fritschi et al. verwendeten den damals zur Verfügung sehenden technischen Möglichkeiten entsprechend einen Ziehungsalgorithmus, bei dem ausgehend von einem Zufallsstartwert die nach Gemeinden strukturierte Population mit fixen Auswahlschritten durchwandert wird. Dieses Verfahren führt zu einer geringeren Variation der Anzahl Treffer in einer Gemeinde mit gegebener Gemeindegröße als ein „echtes“ Zufallsverfahren, was die Verzerrung der durch den Stichprobenplan erzeugten Gemeindegrößenverteilung fast vollständig eliminiert. Der von Fritschi et al. angewendete Algorithmus kann als Verfahren angesehen werden, bei dem quasi nach Gemeindegröße geschichtet gezogen wird. Eine systematische Schichtung nach Gemeindegröße kann durchaus wünschenswert sein. Allerdings sollte man dann gleich ein „echtes“ Schichtungsverfahren verwenden, dessen theoretische Eigenschaften bekannt sind (vgl. auch Abschnitt 4).

3 Ein vereinfachtes „geteiltes“ Verfahren

Zwar ist es, wie oben dargelegt, möglich, das Verfahren von Fritschi et al. (1976) so zu modifizieren, dass unverzerrte Stichproben erzeugt werden. Alternativ könnten auch nachträglich Gewichte berechnet werden, um der Verzerrung bei der Datenanalyse entgegenzuwirken. Ich möchte hier aber zeigen, dass beide Möglichkeiten wenig praktische Relevanz besitzen, da es ein Verfahren zur Ziehung einer „geteilten“ Stichprobe gibt, das in vergleichbarer Art zu einer Reduktion der Anzahl Gemeinden in der Stichprobe führt, jedoch einfacher umzusetzen ist und erst noch etwas bessere statistische Eigenschaften aufweist. Die Komplikationen des Verfahrens von Fritschi et al. rühren daher, dass die Gemeinden *ex post*, also erst *nach* der ersten Ziehung einer einfachen Wahrscheinlichkeitsauswahl aufgrund der empirisch realisierten Anzahl Treffer in zwei Gruppen eingeteilt werden. Eine offensichtliche Strategie zur Vermeidung dieser Probleme ist, die Gruppenzuteilung ganz einfach *ex ante* aufgrund der theoretisch erwarteten Anzahl Treffer vorzunehmen. Ich schlage also ein Verfahren vor, bei dem – gegeben die Trefferwahrscheinlichkeit $p = n/N$ – die Gemeinden zuerst in eine Gruppe mit $E(X_j) = p \cdot N_j \geq k$ und eine Gruppe mit $E(X_j) < k$ aufgeteilt werden. In einem zweiten Schritt wird dann aus der in Gruppe 1 mit den großen Gemeinden enthaltenen Population eine einfache Wahrscheinlichkeitsauswahl gezogen und aus Gruppe 2 mit den kleinen Gemeinden eine Klumpenstrichprobe. Der angestrebte Stichprobenumfang n wird dabei im Verhältnis der durch die beiden Gruppen abgedeckten Bevölkerungsanteile aufgeteilt. Das heißt, der zu realisierende Stichprobenanteil in Gruppe 1 ist

$$n_1 = \frac{n}{N} \cdot \sum_{E(X_j) \geq k} N_j = \sum_{E(X_j) \geq k} E(X_j)$$

und die Anzahl in Gruppe 2 zu ziehender Klumpen beträgt

$$n_k = \frac{n_2}{k} = \frac{n}{k \cdot N} \cdot \sum_{E(X_j) < k} N_j = \frac{1}{k} \sum_{E(X_j) < k} E(X_j)$$

wobei bei nicht ganzzahligen Ergebnissen für n_1 und n_k in der Praxis Zufallsrundungen einzusetzen sind. Es folgt unmittelbar, dass es sich bei der beschriebenen *ex ante* geteilten Zufallsstichprobe um ein Verfahren mit identischen Auswahlwahrscheinlichkeiten handelt, da

$$P_{ij|E(X_j) \geq k} = \frac{n_1}{\sum_{E(X_j) \geq k} N_j} = P_{ij|E(X_j) < k} = \frac{k \cdot n_k}{\sum_{E(X_j) < k} N_j} = \frac{n}{N}$$

Die praktische Durchführung der Ziehung einer *ex ante* geteilten Zufallsstichprobe ist bedeutend einfacher als die Ziehung einer Stichprobe nach dem korrigierten Verfahren von Fritschi et al. (1976), da keine komplizierten Ergebnisse wie Formel (2) benötigt werden. Das Verfahren hat aber noch einen anderen Vorteil: Durch die vorgängige Aufteilung der Population entsteht ein Schichtungseffekt, der die Effizienz der

Stichprobe für Merkmale, die einen Zusammenhang zur Gemeindegröße aufweisen, im Vergleich zum Stichprobenplan von Fritschi et al. zusätzlich erhöht.

Tabelle 2 zeigt die Ergebnisse einer Simulation, bei der das hier vorgeschlagene *ex ante* geteilte Stichprobenverfahren mit einer einfachen Wahrscheinlichkeitsauswahl, einer Klumpenstichprobe und dem *ex post* geteilten Stichprobenverfahren (dem korrigierten Berner Stichprobenplan nach Formel 2) verglichen wird. Grundlage der Simulation ist wiederum das Gemeindeverzeichnis 2002 der Schweiz (Bundesamt für Statistik 2002). Gezogen werden Stichproben im Umfang von $n = 1000$. Die Klumpengröße in den geklumpten Stichproben beträgt $k = 8$.[6] Die Simulationsergebnisse in Tabelle 2 beziehen sich auf die Stichprobenschätzer der Erwartungswerte von verschiedenen Merkmalen mit unterschiedlichem linearen Zusammenhang zur der Gemeindegröße (*r*) und mit unterschiedlicher Intra-Klassen-Korrelation in den Gemeinden (ρ; die Intra-Klassen-Korrelation ist ein Maß für die interne Homogenität in den Gemeinden). Sämtliche Merkmale sind standardisiert, das heißt, sie haben in der Population einen Mittelwert von 0 und eine Standardabweichung von 1. Dargestellt sind für jedes Stichprobenverfahren die Durchschnitte und Standardabweichungen der Mittelwertsschätzer über 10’000 Ziehungen.

[Tabelle 2 etwa hier.]

Im Fuß von Tabelle 2 finden sich einige Angaben zur durchschnittlichen Anzahl PSU (*Primary Sampling Units*, d.h. Anzahl Klumpen)[7] und der durchschnittlichen Anzahl Gemeinden in den Stichproben. Man erkennt sehr schön, dass alle Klumpungsverfahren in vergleichbarer Weise zu einer Reduktion der Anzahl Gemeinden auf etwas mehr als 100 führen – im Vergleich zu den knapp 600 Gemeinden, die bei einer einfachen Wahrscheinlichkeitsauswahl in die Stichprobe gelangen. Verbunden mit der Verringerung der Anzahl Gemeinden ist naturgemäß eine Reduktion der Anzahl PSU, wobei aber diese Reduktion beim Berner Stichprobenplan (275 bzw. 266 PSU) deutlich geringer ausfällt als bei der reinen Klumpenstichprobe (125 PSU). Gerade in dieser geringeren Reduktion der Anzahl PSU bei gleichzeitig vergleichbarer Reduktion der Anzahl Gemeinden liegt die Stärke des Berner Stichprobenplans gegenüber der reinen Klumpenstichprobe, wobei dieser Vorteil allerdings nur unter bestimmten, noch zu identifizierenden Bedingungen tatsächlich zu einer verbesserten Stichprobeneffizienz führt.

An den Resultaten im Hauptteil von Tabelle 2 erkennt man, dass mit allen Stichprobenverfahren die Populationsmittelwerte der vier Merkmale erwartungstreu geschätzt werden: Die Populationsmittelwerte werden mit allen Verfahren im Mittel über die 10’000 Replikationen praktisch exakt reproduziert. Die Verfahren unterscheiden sich aber deutlich hinsichtlich der Streuung der Mittelwertsschätzer über

[6] Man beachte, dass mit diesen Simulationsparametern nur gerade acht der knapp 3000 Gemeinden in den ungeklumpten Zweig der *ex ante* geteilten Stichprobe gelangen (es sind dies Zürich, Genf, Basel, Bern, Lausanne, Winterthur, St. Gallen und Luzern).

[7] Bei der einfachen Wahrscheinlichkeitsauswahl sind die PSU identisch mit den Zielpersonen, das heisst, eine Stichprobe von Umfang *n* enthält *n* Klumpen mit jeweils einem Mitglied. Die reine Klumpenstichprobe enthält *n*/*k* Klumpen mit je *k* Mitgliedern. Bei den geteilten Verfahren enthalten die Stichproben eine Mischung aus Ein-Personen- und *k*-Personen-Klumpen.

die einzelnen Stichproben. Zum Beispiel ist bei der Klumpenstichprobe der Standardfehler des Mittelwertsschätzer für Merkmal X1, welches perfekt mit der Gemeindegröße korreliert (X1 ist eine lineare Transformation der Gemeindegröße), fast dreimal so groß wie bei der einfachen Wahrscheinlichkeitsauswahl. Der Design-Effekt in der Simulation, also das Verhältnis der Varianzen, beträgt in diesem Fall 8.12, was in etwa dem nach der Formel von Kish (1965, S. 162) erwarteten Design-Effekt für Klumpenstichproben von $1+\rho(k-1)=8$ entspricht (vgl. Tabelle 3). Beim (korrigierten) Berner Stichprobenplan ist der Standardfehler für Merkmal X1 jedoch nur geringfügig größer als bei der einfachen Zufallsauswahl und beim *ex ante* geteilten Verfahren ist er sogar bedeutend kleiner (der Design-Effekt beträgt 0.36). Letzteres liegt am angesprochenen Schichtungseffekt durch die vorgängige Teilung der Stichprobe in zwei Gemeindegrößengruppen, aus denen dann separate Stichproben gezogen werden.

[Tabelle 3 etwa hier.]

Bei Merkmal X2, das mittelstark mit der Gemeindegröße korreliert, aber über keine über das durch den linearen Zusammenhang bedingte Ausmaß hinausgehende Intra-Klassen-Korrelation verfügt,[8] schneidet der Berner Stichprobenplan ebenfalls sehr gut ab. Der Design-Effekt ist für die *ex ante* geteilte Stichprobe immer noch Varianz vermindernd und beträgt 0.86 (gegenüber 2.78 für die Klumpenstichprobe; Tabelle 3, Spalte 2). Wird nun die Intra-Klassen-Korrelation aber systematisch erhöht, verlieren die geteilten Stichprobenverfahren zunehmen an Boden. Merkmal X3 weist wie X2 einen linearen Zusammenhang von 0.5 auf, die Intra-Klassen-Korrelation wurde aber künstlich von den minimal 0.25 auf ebenfalls 0.5 angehoben. Die Streuung ist nun in der *ex ante* geteilte Stichprobe deutlich höher als in der einfachen Wahrscheinlichkeitsauswahl, der Design-Effekt ist aber mit 2.44 immer noch nur etwa halb so groß wie bei der reinen Klumpenstichprobe.

Die Vorteile des Berner Stichprobenplans verschwinden schließlich fast vollständig, wenn die interne Homogenität der Gemeinden auf hohem Niveau beibehalten, der lineare Zusammenhang zur Gemeindegröße aber eliminiert wird. Merkmal X4 ist unabhängig von der Gemeindegröße an sich, weist aber eine relativ starke interne Homogenität in den Gemeinden auf. Das heißt, Personen aus der gleichen Gemeinde sind sich bezüglich X4 – unabhängig von der Gemeindegröße – im Durchschnitt bedeutend ähnlicher als Personen aus unterschiedlichen Gemeinden. Die Eigenschaften des Berner Stichprobenplans verschlechtern sich hier verglichen mit der einfachen Wahrscheinlichkeitsauswahl deutlich. Der Design-Effekt erreicht nun fast einen Wert, wie er für die reine Klumpenstichprobe festgestellt werden kann, und liegt – etwas überraschend – deutlich über dem theoretisch erwarteten Design-Effekt für eine reine Klumpenstichprobe mit einer dem Wert für die geteilte Stichprobe entsprechenden Anzahl von 266 Klumpen bzw. PSU (vgl. die letzte Zeile von Tabelle 3). Das heißt, dass sich die größere Anzahl PSU im geteilten Verfahren bei Merkmal X4 nicht voll ausspielen kann. Daraus lässt sich schließen, dass es auf die genaue Struktur ankommt, in der sich die Intra-Klassen-Korrelation eines Merkmals präsentiert. Bei den bisherigen Merkmalen wurde von homoskedastischer Varianz innerhalb der Gemeinden

[8] Aus dem linearen Zusammenhang zwischen einem Merkmal X und der Gemeindegröße folgt ein Mindestmaß an interner Homogenität in den Gemeinden von $\rho_x^{min} = r^2$.

ausgegangen. Die Intra-Klassen-Korrelation kann nun aber wiederum selbst eine Funktion der Gemeindegröße sein, was sich offensichtlich auf die Effizienz des Berner Stichprobenplans auszuwirken scheint. Merkmal X5 repräsentiert den Extremfall in dem in den großen Gemeinden mit $E(X_j) \geq k$ perfekte interne Homogenität herrscht, die kleinen Gemeinden mit $E(X_j) < k$ jedoch eine Intra-Klassen-Korrelation von $\rho = 0$ aufweisen. Die Varianzen in den beiden Teilen der Population wurden zudem so gewählt, dass sich über alle Gemeinden hinweg eine Intra-Klassen-Korrelation von 0.5 ergibt. An den Simulationsresultaten in den Tabellen 2 und 3 erkennt man, dass in dieser Situation, von der man auf den ersten Blick annehmen könnte, sie unterscheide sich kaum von der Situation für Merkmal X4, das *ex ante* geteilte Stichprobenverfahren wiederum sehr gut abschneidet und einen neutralen Design-Effekt von 1 erreicht. An was liegt das? Die Begründung ist klar: In den kleinen Gemeinden ist die geteilte Stichprobe eine Klumpenstichprobe und bei einer Intra-Klassen-Korrelation von 0 ist die Klumpenstichprobe gleich effizient wie eine einfache Wahrscheinlichkeitsauswahl (vgl. unten). In den großen Gemeinden entspricht die geteilte Stichprobe einer einfachen Wahrscheinlichkeitsauswahl, die bei einer Intra-Klassen-Korrelation von 1 einen maximalen Effizienzvorteil gegenüber der Klumpenstichprobe aufweist. Die *ex ante* geteilte Stichprobe *muss* in dieser Situation folglich einen Design-Effekt von 1 haben.[9] Umgekehrt lässt sich auch ableiten, unter welchen Bedingungen die geteilte Stichprobe genau gleich ineffizient ist wie eine reine Klumpenstichprobe. Dies ist der Fall, wenn in den kleinen Gemeinden perfekte interne Homogenität herrscht und die großen Gemeinden perfekte Heterogenität aufweisen. Gegeben die Korrelation zwischen einem Merkmal und der Gemeindegröße ist null, liegt der Berner Stichprobenplan bezüglich der statistischen Effizienz also je nach Struktur der internen Homogenität in den Gemeinden zwischen den beiden Polen der einfachen Wahrscheinlichkeitsauswahl und der Klumpenstichprobe. Bei Korrelation zwischen dem Merkmal und der Gemeindegröße (bzw. genauer: bei positiver Intra-Klassen-Korrelation in den beiden Gemeindegrößenschichten) kommt beim *ex ante* geteilten Stichprobenverfahren zusätzlich der Schichtungseffekt zum Tragen, so dass für einzelne Merkmale unter Umständen eine größere statistische Effizienz als in der einfachen Wahrscheinlichkeitsauswahl erreicht werden kann.

Merkmal X6 illustriert weiterhin den Fall, in dem mit allen Stichprobenverfahren die gleichen Ergebnisse erzielt werden. Wenn keine Intra-Klassen-Korrelation in den Gemeinden besteht, das heißt, wenn sich Personen aus der gleichen Gemeinde bezüglich eines Merkmals nicht ähnlicher sind als Personen aus unterschiedlichen Gemeinden, dann spielt es für die Effizienz der Stichprobe keine Rolle, ob die Stichprobe nach Gemeinden geklumpt ist oder nicht. Selbst in der reinen Klumpenstichprobe ist in diesem (in der Realität wohl eher selten anzutreffenden) Fall also der Standardfehler nicht größer als in der einfachen Wahrscheinlichkeitsauswahl.

Tabelle 3 mit den Design-Effekten enthält schließlich noch den Vergleich zwischen dem *ex ante* geteilten Stichprobenverfahren und dem *ex post* geteilten Verfahren, bei dem die Auswahlwahrscheinlichkeiten nach Formel (2) bestimmt wurden. Man erkennt,

---

[9] Die Tatsache, dass die *ex ante* geteilte Stichprobe geschichtet ist, spielt hier keine Rolle, da X4 unkorreliert ist mit der Gemeindegrösse. Das heisst, der Mittelwert von X4 ist in beiden Schichten identisch und es kann somit kein Schichtungseffekt entstehen.

dass aufgrund des Schichtungseffekts die *ex ante* geteilte Stichproben in fast allen Situationen zumindest geringfügig besser abschneidet als die *ex post* geteilte Stichprobe und nie eine größere Varianz aufweist. Die Wahl des einfacher umzusetzenden *ex ante* geteilten Stichprobenverfahrens ist somit die „dominante" Strategie: man kann sich gegenüber dem *ex post* geteilten Verfahren nur verbessern, nicht aber verschlechtern.

4 Zusammenfassung und Diskussion

Die in der Schweiz am häufigsten eingesetzte Methode zur Erstellung einer Bevölkerungsstichprobe – die Ziehung aus dem Telefonregister – ist vor allem aufgrund mangelnder Abdeckung der Grundgesamtheit nur mit Vorbehalt für Studien, die wissenschaftlichen Qualitätsstandards erfüllen sollen, in Betracht zu ziehen. Der Berner Stichprobenplan, bei dem die Stichprobe über die Einwohnerregister der Gemeinden bestimmt wird, kann hier eine Alternative sein. Zwar sind die Kosten und der benötigte Zeitaufwand der Adressbeschaffung höher als bei einer Telefonregisterstichprobe, die zweifelsohne stark verbesserte Qualität der resultierenden Ausgangsstichprobe lässt den Zusatzaufwand, der sich durch die Klumpung der Stichprobe auf wenige Gemeinden in überschaubarem Rahmen bewegt, aber durchaus als lohnenswert erscheinen. Ohnehin ist der Aufwand für die Stichprobenziehung eher verschwindend, wenn er am Aufwand gemessen wird, der dann in der Regel für die Durchführung der Datenerhebung zu betreiben ist. Man spart sicherlich am falschen Ort, wenn man auf billige, qualitativ schlechte Stichprobenverfahren setzt, auch wenn durch die Kosteneinsparung ein paar zusätzliche Interviews realisieren werden können. Denn was nützt eine größere Fallzahl, wenn die Stichprobe verzerrt ist?

Die Eigenschaften des Berner Stichprobenplans wurden in den Abschnitten 2 und 3 erläutert. Es wurde erstens gezeigt, dass das ursprünglich von Fritschi et al. (1976) vorgeschlagene Verfahren zu leicht verzerrten Stichproben führt, was aber mit ein paar Modifikationen korrigiert werden kann. Zudem wurde mit dem *ex ante* geteilten Stichprobenverfahren ein alternativer Ansatz vorgeschlagen, der einfacher umzusetzen ist und erst noch etwas bessere statistische Eigenschaften aufweist als der ursprüngliche Stichprobenplan von Fritschi et al. (1976). Unabhängig davon, ob man jetzt das Originalverfahren verwendet oder das hier vorgeschlagene vereinfachte Verfahren, zeigt sich der Berner Stichprobenplan vor allem dann gegenüber der Klumpenstichprobe vorteilhaft, wenn ein Zusammenhang besteht zwischen dem Untersuchungsmerkmal und der Gemeindegröße. Bei Merkmalen ohne Zusammenhang zur Gemeindegröße sind die Verhältnisse etwas komplizierter. Die genaue Struktur der Intra-Klassen-Korrelation ist dann von Bedeutung. Etwas überraschend erzielt der Berner Stichprobenplan besonders gute Ergebnisse, wenn die interne Homogenität in den großen Gemeinden am stärksten ist. Dies erscheint etwas ungünstig, da man wohl eher davon ausgehen würde, dass, falls überhaupt Unterschiede bestehen, für die meisten Merkmale die interne Homogenität in kleinen, ländlichen Gemeinden stärker ist als in den Städten. Trotz dieses Vorbehalts zeigen die Simulationen jedoch, dass der Berner Stichprobenplan in den meisten Fällen deutlich besser abschneidet als ein einfaches Klumpungsverfahren und somit bei etwa gleich hohem Aufwand für die Beschaffung der Adressen der Zielpersonen insgesamt zu effizienteren Stichproben führt. Zudem ist nicht auszuschließen, dass sich die Stichprobeneffizienz noch weiter erhöhen lässt, wenn das Berner Stichprobenverfahren mit anderen Techniken wie zum Beispiel einer geschichteten Klumpenstichprobe auf Grundlage von Gemeindetypologien und -größen (vergleiche etwa Buchmann und Sacchi 1997) kombiniert wird. Dies wäre allenfalls in einer Folgestudie zu klären.

Auch gestaltet sich die Datenanalyse beim Berner Stichprobenplan kaum komplizierter als bei einer einfachen Klumpenstichprobe. Bei beiden Verfahren müssen Schätzer verwendet werden, die die Klumpenstruktur der Daten in Rechnung stellen, um die Varianzen der statistischen Koeffizienten nicht zu unterschätzen. Bei Stichproben auf Grundlage des *ex ante* geteilten Verfahrens ist zudem zu berücksichtigen, dass die Stichprobe aus zwei Schichten gezogen wurde, was die Varianzen im Allgemeinen verringert. Komplexe Schätzer, die Klumpungs- und Schichtungseffekte in angemessener Weise behandeln, stehen für eine Vielzahl statistischer Verfahren zur Verfügung (zur Analyse von komplexen Stichproben vgl. z.B. Lohr 1999 oder Levy und Lemeshow 2002) und sind mittlerweile in allgemeinen Statistik-Programmen recht gut unterstützt (vgl. z.B. StataCorp 2005). Surveys, bei denen es weder Schichtungs- und/oder Klumpenstrukturen noch Gewichte zur Korrektur unterschiedlicher Auswahlwahrscheinlichkeiten zu berücksichtigen gilt, sind ohnehin recht selten. Die Datenanalyse dürfte somit beim Berner Stichprobenplan keine zusätzlichen Schwierigkeiten bereiten.

Bei allen Vorteilen des Verfahrens ist eine zentrale Anwendungsvoraussetzung des Berner Stichprobenplanes jedoch noch zu nennen. Die hier präsentierten Simulationsergebnisse beruhen alle auf dem Gemeindeverzeichnis 2002 der Schweiz. Eine Eigenschaft der Schweizer Gemeinden ist, dass die Verteilung der Anzahl Personen, die in diesen Gemeinden leben, sehr schief ist. Das heißt, es gibt hunderte kleiner Gemeinden aber nur eine Hand voll große. Natürlich ist eine solche Datenstruktur eine Bedingung dafür, dass der Berner Stichprobenplan überhaupt zu anderen Resultaten führt als ein einfaches Klumpungsverfahren. Wird das geteilte Stichprobenverfahren auf Aggregateinheiten angewendet, die sich in ihrer Größe nur wenig unterscheiden, ist die Wahrscheinlichkeit groß, dass keine einzige Einheit die notwendige Mindestzahl an Treffern für die Aufnahme in den ungeklumpten Zweig der Stichprobe erreicht. Eine schiefe Verteilung der Größen der verwendeten Aggregateinheiten ist also eine Voraussetzung für eine sinnvolle Anwendung des Berner Stichprobenplans. Dies lässt zum Beispiel den Einsatz des Verfahrens auf Ebene von Stimmbezirken, deren Größen in der Regel nicht so stark schwanken, als weniger günstig erscheinen.

5 Anhang

Die folgenden Ergebnisse werden zur Lösung von Gleichung (2) benötigt. Für eine binomialverteilte Zufallsvariable $X \sim B(n,p)$ ist die Wahrscheinlichkeit, k oder mehr Treffer zu erreichen, bekanntlich gegeben als

$$P(X \geq k) = \sum_{x=k}^{n} \binom{n}{x} p^x (1-p)^{n-x}$$

(vgl. z.B. Evans et al. 2000, S. 43-47). Funktionen zur automatischen Berechnung von $P(X \geq k)$ stehen in den meisten Statistikprogrammen zur Verfügung. Etwas aufwändiger ist die Berechnung von $E(X \mid X \geq k)$, also dem Erwartungswert einer binomialverteilten Zufallsvariable, gegeben eine bestimmte Mindestzahl an Treffern wird erreicht. Der Erwartungswert einer diskreten Variablen entspricht dem Mittel der mit den Auftretenswahrscheinlichkeiten gewichteten Ausprägungen, im vorliegenden Fall also

$$E(X \mid X \geq k) = \frac{k \cdot P(X = k) + (k+1) \cdot P(X = k+1) + \cdots + n \cdot P(X = n)}{P(X \geq k)}$$

was sich vereinfachen lässt zu

$$E(X \mid X \geq k) = k + \frac{P(X \geq k+1) + \cdots + P(X = n)}{P(X \geq k)}$$

Zur zeitsparenden Berechnung von $E(X \mid X \geq k)$ für $k < n/2$ sind zudem die folgenden Ergebnisse von Bedeutung:

$$E(X) = P(X \geq k) \cdot E(X \mid X \geq k) + P(X < k) \cdot E(X \mid X < k)$$

$$E(X \mid X < k) = k - 1 - \frac{P(X < 1) + \cdots + P(X < k-1)}{P(X < k)}$$

6 Literaturverzeichnis

Babbie, E. R. (1979): The Practice of Social Research. Second Edition. Belmont, CA: Wadsworth.

Buchmann, M.; Sacchi, S. (1997). Berufsverlauf und Berufsidentität im sozio-technischen Wandel. Konzeption, Methodik und Repräsentativität einer retrospektiven Befragung der Geburtsjahrgänge 1949-51 und 1959-61. ETH Zürich.

Bundesamt für Statistik (2002): Gemeinde- und Ortschaftenverzeichnis 2002. Neuchâtel: BFS.

Evans, M.; Hastings, N.; Peacock, B. (2000): Statistical Distributions. Third Edition. New York: Wiley.

Fritschi, P.; Meyer, R.; Schweizer, W. (1976): Ein neuer Stichprobenplan für ein gesamtschweizerisches Sample. In: Schweizerische Zeitschrift für Soziologie 2(3), S. 149–158.

Gabler, S.; Häder, S. (1997): Überlegungen zu einem Stichprobendesign für Telefonumfragen in Deutschland. ZUMA-Nachrichten 41, S. 7-19.

Häder, S.; Glemser, A. (2006): Stichprobenziehung für Telefonumfragen in Deutschland. In: Diekmann, A. (Hrsg). Methoden der Sozialforschung. Sonderheft 44 der Kölner Zeitschrift für Soziologie und Sozialpsychologie. Wiesbaden: VS Verlag, S. 148-171.

Handcock, M. S.; Morris, M. (1999): Relative Distribution Methods in the Social Sciences. New York: Springer.

Jann, B. (2001): Stichprobenziehung aus TwixTel. Universität Bern. Erhältlich unter http://www.socio.ethz.ch/people/jannb/stuff/stichprobenziehung2up.pdf.

Kish, L. (1965): Survey Sampling. New York: Wiley.

Leu, R. E.; Burri, S.; Prister, T. (1997): Lebensqualität und Armut in der Schweiz. Bern: Haupt.

Levy, P. S.; Lemeshow, S. (1999): Sampling of Populations: Methods and Applications, 3rd ed. New York: Wiley.

Lohr, S. L. (1999): Sampling: Design and Analysis. Pacific Grove: Duxbury Press.

Meyer R.; Haltiner, K.; Hofer, R.; Iff, H.; Rüegg, W. (1982): Fragen an die Zukunft. Die Bedeutung von Beruf, Bildung und Politik für die zwanzigjährigen Schweizerinnen und Schweizer. Aarau und Frankfurt am Main: Sauerländer.

Sasieni, P. (1995): Symmetric nearest neighbor linear smoothers. In: Stata Technical Bulletin 24, S. 10-14.

Schmugge, S.; Grau, P. (1998): Telefon-Anschlüsse der privaten Haushalte in der Schweiz. Situationsanalyse 1998. Luzern: LINK Institut.

Schmugge, S.; Grau, P. (2000): Sind die SchweizerInnen überhaupt noch zu erreichen? Telefonanschlüsse der privaten Haushalte in der Schweiz im Jahr 2000. Luzern: LINK Institut.

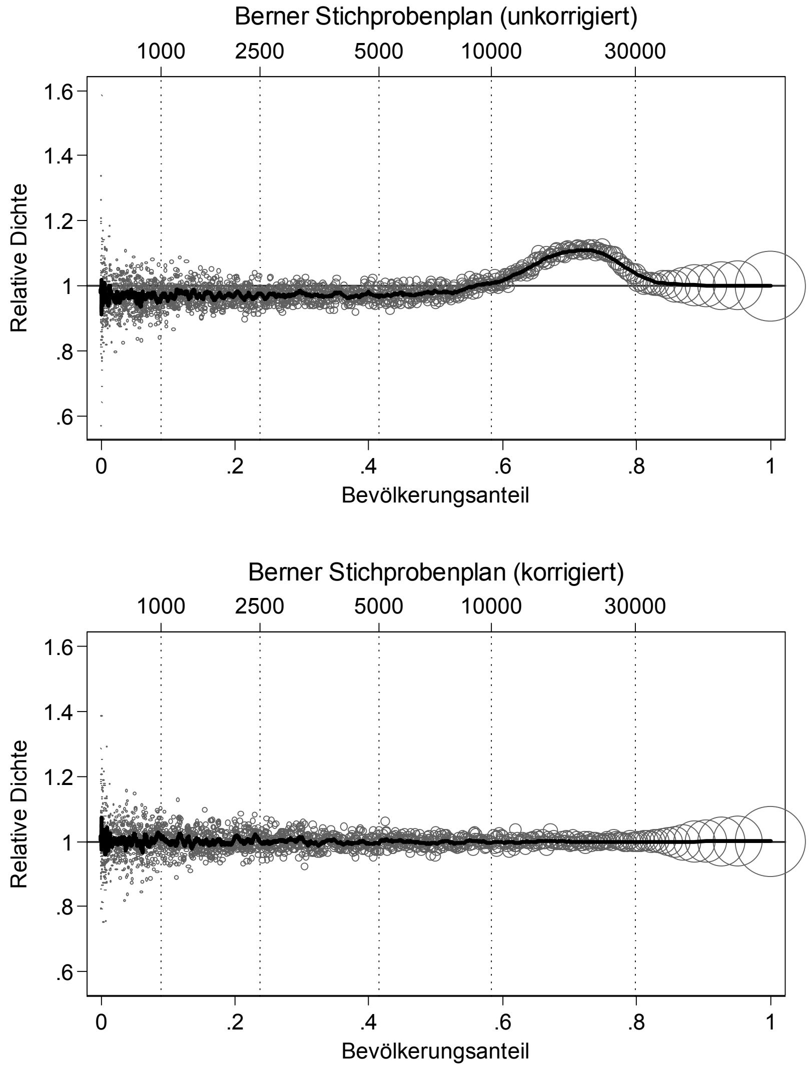
Schnell, R. (1991): Wer ist das Volk? Zur faktischen Grundgesamtheit bei „allgemeinen Bevölkerungsumfragen“: Undercoverage, Schwererreichbare und Nichtbefragbare. In: Kölner Zeitschrift für Soziologie und Sozialpsychologie 43(1), S. 106-137.

Schnell, R. (1997): Praktische Ziehung von Zufallsstichproben für Telefon-Surveys. In: ZA-Informationen 40, S. 45-59.

StataCorp (2005): Stata Survey Data Reference Manual. College Station, Texas: Stata Press.

Wydler, H.; Walter, T.; Hattich, A.; Hornung, R.; Gutzwiller, F. (1996): Die Gesundheit 20jähriger in der Schweiz: Ergebnisse der PRP 1993. Aarau: Sauerländer.

Abbildung 1: Vergleich der Gemeindeverteilung des Berner Stichprobenplans (gemittelt über 10’000 Ziehungen) mit der Gemeindeverteilung in der Grundgesamtheit



Legende: Dargestellt ist für jede Gemeinde das Verhältnis zwischen ihrem durchschnittlichen Anteil in der Stichprobe (10’000 Ziehungen mit Stichprobengröße 3750 und Klumpengröße 10) und ihrem Anteil an der Gesamtbevölkerung. Ein Wert von eins bedeutet, dass die Mitglieder der Gemeinde im Durchschnitt genau ihrem Anteil an der Gesamtbevölkerung entsprechend repräsentiert werden (zum Konzept der relativen Verteilung siehe Handcock und Morris 1999). Die Gemeinden sind ihrer Größe nach auf der proportional zum kumulierten Bevölkerungsanteil skalierten X-Achse geordnet, zudem richtet sich die Größe der Symbole nach der Anzahl Einwohner in den entsprechenden Gemeinden. Da in einer endlichen Simulation die Werte vor allem für kleine Gemeinden beträchtlich streuen, wurde zur Verdeutlichung des Trends eine geglättete Kurve eingezeichnet (Running-Line-Methode mit jeweils beidseitig 10 benachbarten Punkten gewichtet nach Gemeindegröße, vgl. Sasieni 1995).

Tabelle 1: Gemeindegrößenverteilung im Berner Stichprobenplan

Gemeindegröße	Bevölkerungs-anteil	Stichprobenverteilungen (10’000 Replikationen)		
		Berner Stich-probenplan (unkorrigiert)	Berner Stich-probenplan (korrigiert)	einfache Zufallsauswahl
1 – 249	0.98	0.94	0.98	0.98
250 – 499	2.47	2.39	2.48	2.47
500 – 999	5.56	5.36	5.58	5.57
1000 – 1999	10.49	10.18	10.48	10.48
2000 – 3499	12.06	11.71	12.08	12.07
3500 – 4999	9.89	9.60	9.87	9.90
5000 – 9999	16.79	16.49	16.77	16.80
10000 – 49999	25.67	27.25	25.66	25.65
50000 und mehr	16.09	16.08	16.10	16.08
Total	100.00	100.00	100.00	100.00

Quelle: Simulationsergebnisse (10’000 Replikationen) auf Grundlage des Gemeindeverzeichnisses 2002 der Schweiz (BFS 2002, 2876 Gemeinden mit insgesamt 7’241’468 Einwohnern) mit Ziel-Stichprobengröße 3750 und Klumpengröße 10.

Tabelle 2: Mittelwerte und Standardfehler (in Klammern) von Erwartungswertsschätzern verschiedener Stichprobenverfahren

	Erwartungs-wert	Simulationsergebnisse (10’000 Replikationen)			
		einfache Zufalls-auswahl	Klumpen-stichprobe	Berner Stichprobenplan	
				ex post geteilte Stichprobe	*ex ante* geteilte Stichprobe
X1 ($r = 1.00$, $\rho = 1.00$)	0.000	– 0.000	0.001	0.000	0.000
		(0.032)	(0.090)	(0.033)	(0.019)
X2 ($r = 0.50$, $\rho = 0.25$)	0.000	– 0.000	0.000	0.000	– 0.000
		(0.031)	(0.052)	(0.032)	(0.029)
X3 ($r = 0.50$, $\rho = 0.50$)	0.000	– 0.000	– 0.000	0.000	– 0.001
		(0.031)	(0.067)	(0.051)	(0.049)
X4 ($r = 0.00$, $\rho = 0.50$)	0.000	0.000	– 0.001	– 0.000	– 0.001
		(0.031)	(0.067)	(0.064)	(0.065)
X5 ($r = 0.00$, $\rho = 0.50^{a}$)	0.000	– 0.000	– 0.001	0.000	0.000
		(0.032)	(0.068)	(0.033)	(0.032)
X6 ($r = 0.00$, $\rho = 0.00$)	0.000	– 0.000	– 0.000	0.000	– 0.000
		(0.032)	(0.032)	(0.032)	(0.032)
Anzahl PSU		1000.0	125.0	275.1	265.8
Anzahl Gemeinden		568.8	104.7	105.9	105.9
Fallzahl		1000.0	1000.0	1000.0	1000.0

Quelle: Simulationsergebnisse (10’000 Replikationen) auf Grundlage des Gemeindeverzeichnisses 2002 der Schweiz (BFS 2002, 2876 Gemeinden mit insgesamt 7’241’468 Einwohnern) mit Ziel-Stichprobengröße 1000 und Klumpengröße 8. Bei den Variablen X1 bis X6 handelt es sich um künstlich für die Population erzeugte Merkmale (*r*: Korrelation mit der Gemeindegröße; ρ: Intra-Klassen-Korrelation, [a] mit $\rho = 1$ in den großen und $\rho = 0$ in den kleinen Gemeinden).

Tabelle 3: Simulierte und theoretische Design-Effekte (Verhältnis der Varianzen)

	X1 ($r = 1.00$, $\rho = 1.00$)	X2 ($r = 0.50$, $\rho = 0.25$)	X3 ($r = 0.50$, $\rho = 0.50$)	X4 ($r = 0.00$, $\rho = 0.50$)	X5 ($r = 0.00$, $\rho = 0.50$ [a])	X6 ($r = 0.00$, $\rho = 0.00$)
Klumpenstichprobe vs. einfache Zufallsauswahl	8.12	2.78	4.61	4.60	4.57	1.00
ex ante geteilte Stichprobe vs. einfache Zufallsauswahl	0.36	0.86	2.44	4.22	1.00	1.01
ex ante geteilte Stichprobe vs. *ex post* geteilte Stichprobe	0.33	0.83	0.92	1.00	0.93	1.01
Theoretischer Design-Effekt für eine Klumpenstichprobe mit 125 Klumpen	8.00	2.75	4.50	4.50	4.50	1.00
Theoretischer Design-Effekt für eine Klumpenstichprobe mit 266 Klumpen	3.76	1.69	2.38	2.38	2.38	1.00

Quelle: Beruhend auf den Zahlen in Tabelle 2. r: Korrelation mit der Gemeindegröße; ρ: Intra-Klassen-Korrelation, [a] mit $\rho = 1$ in den großen und $\rho = 0$ in den kleinen Gemeinden